

1646

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/545,772

DATE: 02/02/2001
 TIME: 16:23:43

Input Set : A:\42052200.txt
 Output Set: N:\CRF3\02022001\I545772.raw

ENTERED

RECEIVED

FEB 14 2001

TECH CENTER 1600/2900

3 <110> APPLICANT: WILKINS, TRACEY D.
 4 LYERLY, DAVID M.
 5 MONCRIEF, J. SCOTT
 6 PAVLIAKOVA, DANKA
 7 SCHNEERSON, RACHEL
 8 ROBBINS, JOHN B.
 10 <120> TITLE OF INVENTION: RECOMBINANT TOXIN A PROTEIN CARRIER FOR POLYSACCHARIDE
 11 CONJUGATE VACCINES
 13 <130> FILE REFERENCE: 420522000100
 15 <140> CURRENT APPLICATION NUMBER: 09/545,772
 C--> 16 <141> CURRENT FILING DATE: 2000-04-10
 18 <160> NUMBER OF SEQ ID NOS: 2
 20 <170> SOFTWARE: PatentIn Ver. 2.1
 22 <210> SEQ ID NO: 1
 23 <211> LENGTH: 2507
 24 <212> TYPE: DNA
 25 <213> ORGANISM: Clostridium difficile
 27 <400> SEQUENCE: 1
 28 gatccctatag aatttcaactt agtaactgga tggcaaaacta tcaatggttaa aaaatattat 60
 29 ttgatataaa atactggagc agctttaact agttataaaa ttattaatgg taaacacttt 120
 30 tattttaata atgatgtgtg gatgcagttg ggagtattta aaggacctga tggatttgaa 180
 31 tattttgcac ctgccaatat tcaaaaataat aacatagaag gtcaggctat agtttatcaa 240
 32 agtaaatctt taactttgaa tggcaaaaaa tattattttg ataataactc aaaagcagtc 300
 33 actggatgga gaattattta caatgagaaa tattacttta atcctaataa tgcatttgct 360
 34 gcaatcggaat tgcgaagtaat tgacaataat aagtattatt tcaatcctga cactgctatc 420
 35 atctcaaaag gttggcgagc tgttaaatggt agtagatact actttgatac tgataccgct 480
 36 attgacctta atgggtataa aactattgat ggtaaacact tttattttga tagtgattgt 540
 37 gtatgtaaaa taggtgtgtt tagtacctct aatggatttg aatatlttgc acctgctaatt 600
 38 acttataata ataacataga aggtcaggct atagtttatc aaagtaaat cttaaactttg 660
 39 aatggtaaaa aatattactt tgataataac tcaaaagcag ttaccggatg gcaaaactatt 720
 40 gatagtataa aatattactt taataactac actgctgaag cagctactgg atggcaaaact 780
 41 attgatggta aaaaatatta ctttaatact aacactgctg aagcagctac tggatggcaa 840
 42 actatlgatg gtaaaaaata ttactttaat actaacactg ctatagcttc aactggttat 900
 43 acaattatta atggtaaaac tttttatttt aatactgatg gtattatgca gataggagt 960
 44 tttaaaggac ctaatggatt tgaatatttt gcaactgcta atacggatgc taacaacata 1020
 45 gaaggtaaac ctatacttta ccaaaatgaa ttcttaactt tgaatggtaa aaaatattac 1080
 46 tttggtagtg actcaaaagc agttactgga tggagaatta ttaacaataa gaaatattac 1140
 47 tttaaacctc ataattgctat tgcctgaatt catctatgca ctataaataa tgacaagtat 1200
 48 tacttttagt atgatggaat tcttcaaaat ggatatatta ctattgaaag aaataatttc 1260
 49 latltlgatg ctaataatga atctaaaatg gtaacaggag latttaaagg acctaatgga 1320
 50 tttagatatt ttgcacctgc taatactcac aataataaca tagaagggtcaggctatagtt 1380
 51 taccagaaca aattcttaac ttggaatggc aaaaaataat attttgataa tgactcaaaa 1440
 52 gcaatctact gatggcaaac cattgatggc aaaaaatatt actttaatct taacactgct 1500
 53 gaagcagcta ctggatggca aactattgat ggtaaaaaat attactttta tcttaacact 1560
 54 gctgaagcag ctactggatg gcaaaactat gatggtaaaa aalattactt taatactaac 1620
 55 actttcatag cctcaactgg ttatacaagt attaatggta aacattttta ttttaatact 1680
 56 gatggattta tgcagatagg agtgtttaaa ggacctaatg gatttgaata ctttgcacct 1740

RAW SEQUENCE LISTING

DATE: 02/02/2001

PATENT APPLICATION: US/09/545,772

TIME: 16:23:43

Input Set : A:\42052200.txt

Output Set: N:\CRF3\02022001\I545772.raw

```

57 gctaatacgg atgctaacaa catagaaggt caagctatac tttaacaaaa taaattctta 1800
58 actttgaatg gtaaaaaata ttacttttgg agtgactcaa aagcagltac cggactgcga 1860
59 actattqatg gtaaaaaata ttactttaat actaacactg ctgttgcagt tactggatgg 1920
60 caaactatta atggttaaaaa atactacttt aataactaaca cttctatagc ttcaactggt 1980
61 tatacaatta ttagtggtaa acatttttat tttaatactg atgggtattat gcagatagga 2040
62 gtgtttaaag gacctgatgg atttgaatac ttgacacctg ctaatacaga tgctaacaat 2100
63 atagaaggtc aagctatacg ttatcaaaat agatttcctat atttacatga caatatatat 2160
64 tatttttggt ataatcaaaa agcggctact ggttgggtaa ctattgatgg taatagatat 2220
65 tacttcgagc ctaatacagc tatgggtgcy aatgggttata aaactattga taataaaaaat 2280
66 ttttacttta gaaatggttt acctcagata ggagtgttta aaggggtctaa tggatttgaa 2340
67 tactttgcac ctgctaatac ggaatgtaac aatatagaag gtcaagctat acgttatcaa 2400
68 aatagattcc tacatttact tggaaaaata tattactttg gtaataattc aaaagcagtt 2460
69 actggatggc aaactattaa tggtaaaagta tattacttta tgacctga 2507
72 <210> SEQ ID NO: 2
73 <211> LENGTH: 866
74 <212> TYPE: PRT
75 <213> ORGANISM: Clostridium difficile
77 <400> SEQUENCE: 2
78 Asp Pro Ile Glu Phe Asn Leu Val Thr Gly Trp Gln Thr Ile Asn Gly
79 1 5 10 15
81 Lys Lys Tyr Tyr Phe Asp Ile Asn Thr Gly Ala Ala Leu Thr Ser Tyr
82 20 25 30
84 Lys Ile Ile Asn Gly Lys His Phe Tyr Phe Asn Asn Asp Gly Val Met
85 35 40 45
87 Gln Leu Gly Val Phe Lys Gly Pro Asp Gly Phe Glu Tyr Phe Ala Pro
88 50 55 60
90 Ala Asn Thr Gln Asn Asn Ile Glu Gly Gln Ala Ile Val Tyr Gln
91 65 70 75 80
93 Ser Lys Phe Leu Thr Leu Asn Gly Lys Lys Tyr Tyr Phe Asp Asn Asn
94 85 90 95
96 Ser Lys Ala Val Thr Gly Trp Arg Ile Ile Asn Asn Glu Lys Tyr Tyr
97 100 105 110
99 Phe Asn Pro Asn Asn Ala Ile Ala Ala Val Gly Leu Gln Val Ile Asp
100 115 120 125
102 Asn Asn Lys Tyr Tyr Phe Asn Pro Asp Thr Ala Ile Ile Ser Lys Gly
103 130 135 140
105 Trp Gln Thr Val Asn Gly Ser Arg Tyr Tyr Phe Asp Thr Asp Thr Ala
106 145 150 155 160
108 Ile Ala Phe Asn Gly Tyr Lys Thr Ile Asp Gly Lys His Phe Tyr Phe
109 165 170 175
111 Asp Ser Asp Cys Val Val Lys Ile Gly Val Phe Ser Thr Ser Asn Gly
112 180 185 190
114 Phe Glu Tyr Phe Ala Pro Ala Asn Thr Tyr Asn Asn Asn Ile Glu Gly
115 195 200 205
117 Gln Ala Ile Val Tyr Gln Ser Lys Phe Leu Thr Leu Asn Gly Lys Lys
118 210 215 220
120 Tyr Tyr Phe Asp Asn Asn Ser Lys Ala Val Thr Gly Trp Gln Thr Ile
121 225 230 235 240
123 Asp Ser Lys Lys Tyr Tyr Phe Asn Thr Asn Thr Ala Glu Ala Ala Thr

```

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/545,772
 DATE: 02/02/2001
 TIME: 16:23:43

Input Set : A:\42052200.txt
 Output Set: N:\CRF3\02022001\I545772.raw

124		245		250		255
126	Gly Trp Gln Thr Ile Asp Gly Lys Lys Tyr Tyr Phe Asn Thr Asn Thr					
127		260		265		270
129	Ala Glu Ala Ala Thr Gly Trp Gln Thr Ile Asp Gly Lys Lys Tyr Tyr					
130		275		280		285
132	Phe Asn Thr Asn Thr Ala Ile Ala Ser Thr Gly Tyr Thr Ile Ile Asn					
133		290		295		300
135	Gly Lys His Phe Tyr Phe Asn Thr Asp Gly Ile Met Gln Ile Gly Val					
136	305		310		315	320
138	Phe Lys Gly Pro Asn Gly Phe Glu Tyr Phe Ala Pro Ala Asn Thr Asp					
139		325		330		335
141	Ala Asn Asn Ile Glu Gly Gln Ala Ile Leu Tyr Gln Asn Glu Phe Leu					
142		340		345		350
144	Thr Leu Asn Gly Lys Lys Tyr Tyr Phe Gly Ser Asp Ser Lys Ala Val					
145		355		360		365
147	Thr Gly Trp Arg Ile Ile Asn Asn Lys Lys Tyr Tyr Phe Asn Pro Asn					
148		370		375		380
150	Asn Ala Ile Ala Ala Ile His Leu Cys Thr Ile Asn Asn Asp Lys Tyr					
151	385		390		395	400
153	Tyr Phe Ser Tyr Asp Gly Ile Leu Gln Asn Gly Tyr Ile Thr Ile Glu					
154		405		410		415
156	Arg Asn Asn Phe Tyr Phe Asp Ala Asn Asn Glu Ser Lys Met Val Thr					
157		420		425		430
159	Gly Val Phe Lys Gly Pro Asn Gly Phe Glu Tyr Phe Ala Pro Ala Asn					
160		435		440		445
162	Thr His Asn Asn Asn Ile Glu Gly Gln Ala Ile Val Tyr Gln Asn Lys					
163		450		455		460
165	Phe Leu Thr Leu Asn Gly Lys Lys Tyr Tyr Phe Asp Asn Asp Ser Lys					
166	465		470		475	480
168	Ala Val Thr Gly Trp Gln Thr Ile Asp Gly Lys Lys Tyr Tyr Phe Asn					
169		485		490		495
171	Leu Asn Thr Ala Glu Ala Ala Thr Gly Trp Gln Thr Ile Asp Gly Lys					
172		500		505		510
174	Lys Tyr Tyr Phe Asn Leu Asn Thr Ala Glu Ala Ala Thr Gly Trp Gln					
175		515		520		525
177	Thr Ile Asp Gly Lys Lys Tyr Tyr Phe Asn Thr Asn Thr Phe Ile Ala					
178		530		535		540
180	Ser Thr Gly Tyr Thr Ser Ile Asn Gly Lys His Phe Tyr Phe Asn Thr					
181	545		550		555	560
183	Asp Gly Ile Met Gln Ile Gly Val Phe Lys Gly Pro Asn Gly Phe Glu					
184		565		570		575
186	Tyr Phe Ala Pro Ala Asn Thr Asp Ala Asn Asn Ile Glu Gly Gln Ala					
187		580		585		590
189	Ile Leu Tyr Gln Asn Lys Phe Leu Thr Leu Asn Gly Lys Lys Tyr Tyr					
190		595		600		605
192	Phe Gly Ser Asp Ser Lys Ala Val Thr Gly Leu Arg Thr Ile Asp Gly					
193		610		615		620
195	Lys Lys Tyr Tyr Phe Asn Thr Asn Thr Ala Val Ala Val Thr Gly Trp					
196	625		630		635	640

RAW SEQUENCE LISTING

DATE: 02/02/2001

PATENT APPLICATION: US/09/545,772

TIME: 16:23:43

Input Set : A:\42052200.txt

Output Set: N:\CRF3\02022001\I545772.raw

RECEIVED

FEB 14 2001

TECH CENTER 1600/2900

```

198 Gln Thr Ile Asn Gly Lys Lys Tyr Tyr Phe Asn Thr Asn Thr Ser Ile
199          645          650          655
201 Ala Ser Thr Gly Tyr Thr Ile Ile Ser Gly Lys His Phe Tyr Phe Asn
202          660          665          670
204 Thr Asp Gly Ile Met Gln Ile Gly Val Phe Lys Gly Pro Asp Gly Phe
205          675          680          685
207 Glu Tyr Phe Ala Pro Ala Asn Thr Asp Ala Asn Asn Ile Glu Gly Gln
208          690          695          700
210 Ala Ile Arg Tyr Gln Asn Arg Phe Leu Tyr Leu His Asp Asn Ile Tyr
211 705          710          715          720
213 Tyr Phe Gly Asn Asn Ser Lys Ala Ala Thr Gly Trp Val Thr Ile Asp
214          725          730          735
216 Gly Asn Arg Tyr Tyr Phe Glu Pro Asn Thr Ala Met Gly Ala Asn Gly
217          740          745          750
219 Tyr Lys Thr Ile Asp Asn Lys Asn Phe Tyr Phe Arg Asn Gly Leu Pro
220          755          760          765
222 Gln Ile Gly Val Phe Lys Gly Ser Asn Gly Phe Glu Tyr Phe Ala Pro
223          770          775          780
225 Ala Asn Thr Asp Ala Asn Asn Ile Glu Gln Ala Ile Arg Tyr Gln Asn
226 785          790          795          800
228 Arg Phe Leu His Leu Leu Gly Lys Ile Tyr Tyr Phe Gly Asn Asn Ser
229          805          810          815
231 Lys Ala Val Thr Gly Gly Trp Gln Thr Ile Asn Gly Lys Val Tyr Tyr
232          820          825          830
234 Phe Met Pro Asp Thr Ala Met Ala Ala Ala Gly Gly Leu Phe Glu Asp
235          835          840          845
237 Gly Val Ile Tyr Phe Phe Gly Val Asp Gly Val Lys Ala Pro Gly Ile
238          850          855          860
240 Tyr Gly
241 865

```

VERIFICATION SUMMARY DATE: 02/02/2001
PATENT APPLICATION: US/09/545,772 TIME: 16:23:44

Input Set : A:\42052200.txt
Output Set: N:\CRF3\02022001\I545772.raw

L:16 M:271 C: Current Filing Date differs, Replaced Current Filing Date